

Chapter 5

Speaker Intentions and Intentionality

Michael Haugh and Kasia M. Jaszczolt

1. Introduction

Speaker intentions made their way into contemporary pragmatics through three different but interrelated routes. One of them dates back at least to medieval philosophy and the inquiries into the logic of modal contexts, leading to the study of intentionality. Another begins with ordinary language philosophy of the mid-1950s and the attempts to define meaning through language use, which in turn led to the employment of the concept of speaker's intended effect of an act of communication. The third, and arguably the most influential route, was that of the attempts to rescue formal semantic analyses by employing the concept of meaning which would incorporate not only the truth-conditional content but also the intended implicated messages, forming the overall concept of communicated content. This introduction to intentionality and intentions is structured as follows. In Section 2, we present the philosophical idea of intentionality and explain the relation between intentionality of mental states and linguistic intentionality conveyed through acts of communication. In Section 3 we move to the role of intentions in communication, focusing on the second and third routes mentioned above, starting with Grice's and Searle's views, attempting also a typology, explanation and exemplification of various kinds of speaker intentions distinguished in the literature. Finally, we discuss the question as to where intentions are located, contrasting cognitive, interactional, and discursive perspectives on this issue, and conclude in Section 4 with a brief assessment of the advantages and weaknesses of utilising intentions and intentionality in pragmatic theory.

The first question to ask is why linguists would want to appeal to a concept as murky as intentions and award it the status of an explanandum in pragmatic theory. The main reason is that the possession of the concept of intentional action is crucial for understanding human behaviour. As Anscombe (1957: 83) says in her seminal book *Intention*,

there are many descriptions of happenings which are directly dependent on our possessing the *form* of description of intentional actions. It is easy not to notice this, because it is perfectly possible for some of these descriptions to be of what is done unintentionally. For example ‘offending someone’; one can do this unintentionally, but there would be no such thing if it were never the description of an intentional action.

The notions of intention and intentionality have since been deployed in a multitude of ways in explaining speaker meaning (Grice 1957, 1989), speech acts (Searle 1969, 1983), the development of language (Searle 2007), social development (Tomasello *et al.* 2005), and the cognitive processes underlying action and meaning interpretation (Bara 2010; Pacherie 2006, 2008), to name just a few. However, this proliferation has generated challenges for the conceptual and theoretical status of intentionality and intention in pragmatics, as we will discuss.

2. Intentionality

Intentionality has a long and celebrated tradition in philosophy. Coming from the Latin term *intendere*, meaning aiming in a certain direction, directing thoughts to something, on the analogy to drawing a bow at a target, it has been used to name the property of minds of having content, *aboutness*, being *about* something (Duranti 1999; Harland 1993; Lyons 1995; Jacob 2003; Nuyts 2000; Smith 2008; Jaszczolt 1999; Woodfield 1994). In other words, it means the ability of minds to *represent* objects, properties, or states of affairs. In medieval philosophy, forms of beings consisted of so-called *esse naturale*, or natural objects, and *esse intentionale*, or concepts, mental images, or thoughts. It is the latter that we are interested in here.

The very idea of intentionality dates back to ancient philosophy, as far back as fifth century B.C.E. and Parmenides of Elea. It was taken up by Aristotle and the Stoics (Caston 2007), and was subsequently extensively used in medieval doctrines of knowledge and revived in nineteenth-century phenomenology by Brentano (1874) and Husserl (1900-1901), and later by Meinong, Twardowski, Heidegger, Sartre, Merleau-Ponty and many others. By phenomenology we mean the study of the way in which things (*phenomena*) are presented in consciousness, or generally the study of forms of conscious experience from the first-person

point of view. Mental attitudes such as belief, desire, or want are intentional in that they are *about* something, they have an object. For phenomenologists, things exist as physical objects, but they also have an *intentional* existence, so to speak, in acts of consciousness. Their intentional existence is revealed in our mental states or acts, as for example in (1) or (2).

(1) I am thinking *about my holiday in Australia*.

(2) I hope *to meet Peter Carey one day*.

In the case of hallucinations, the existence is only intentional.

Brentano and Husserl developed intricate arguments concerning the meaning of such mental existence, in particular the question as to whether there are intentional objects that are internal to the thinker's mind. This discussion, albeit fundamental in the study of phenomenology, can only be sketched here. For Brentano, objects of conscious mental acts had the status of mental entities. On this view, the act does not consist of a relation between its subject and an intersubjectively identifiable object but rather can be spelled out in a so-called *adverbial theory*: 'Tom sees a horse' amounts to a property of Tom's 'seeing horsely', so to speak (see Smith and Smith 1995). However, the adverbial theory comes with a considerable limitation. It does not provide for the fact that our acts of consciousness are about things in the world, real world properties, or states of affairs. It was therefore subsequently replaced with the relational theory by Twardowski and Meinong. In the next logical step, Husserl rejected mental objects *tout court*. Current theories of intentionality adopt such a relational view.

Now, this is not to say that where the real object does not exist, 'made-up' substitutes are never needed in a theory of natural language meaning. To give one pertinent example, Clapp (2009) discusses so-called negative existentials, i.e. sentences of the type in (3) and (4) from Clapp (2009: 1422).

(3) The Loch Ness monster does not exist.

(4) Nessie does not exist.

As he puts it, the problem with negative existentials is that they presuppose existence and immediately deny it. We can, of course, stipulate that there is something real to which one refers in (3) and (4), à la Meinong, or say that there is no existential presupposition there, à la Russell. On the modern version of the classical phenomenology, mental objects are

intentional, mind-independent entities; following McGinn (2000), we can introduce the theoretical construct of representation-dependent, intentional objects. Then we replace the dichotomy 'existent – non-existent' with a new dichotomy 'existent – intentional' and thereby vindicate the latter as a theoretical construct.

Intentional content has been the subject of discussions between so-called neo-Fregeans and neo-Russelians (Recanati 1993; Siewert 2006). In the Anglo-American tradition in semantics and pragmatics we often think of the analytic tradition as separate from continental European psychologism with its emphasis on consciousness, mental states, first-person psychology, and the like. However, it is worth remembering that the most groundbreaking achievements in phenomenology and in analytic philosophy took place at the same time, at the end of the nineteenth century and the beginning of the twentieth century and that they are much more interrelated than they are usually taken to be. Frege's (1892) notion of sense, whose offshoots have been so influential in current semantic theory of the Anglo-American tradition, is a direct predecessor of Husserl's phenomenological theory of meaning and object. We briefly attend to this interrelation in what follows.

Husserl held various consecutive ideas of meaning, culminating in his view from *Ideen* (1913) according to which meaning is contained in the objective content of consciousness, called *noema*. An act of consciousness is directed at an object determined by noemata. Here Frege's concept of *sense* can be seen as a clear precursor for Husserl's idea of meaning as distinguished from objects, and hence subsequently also his notion of noema. However, their respective distinctions between meaning and object, and sense and reference, led them in very different directions. Whereas Husserl and other phenomenologists concentrated on experience, Frege and analytic philosophers concentrated on developing a theory of meaning making use of intersubjective generalizations over 'meaning-giving mental acts' such as Frege's sense.

But this departure from psychologism does not reflect a sudden split. It happened gradually through various reanalyses of intentionality and subsequently ascribing intentionality to linguistic acts. So, the question to ask is how the theory of meaning became dissociated from Husserlian meaning as noema, that is from meaning attached to mental acts discussed above. First of all, for Husserl from his *Ideen* period (1913), it was language, rather than a particular mental act of a language user as it was claimed in earlier phenomenological accounts, that was a carrier of meaning. Already in Husserl's work of this period meaning

acquires the status of an abstraction, noematic meaning, attached to relevant mental acts. There is one step from there to freeing theory of meaning from the constraints of psychology. And this step was taken by Frege and his battle against the ‘corrupting intrusion’ of psychology in logic and mathematics, and thereby subsequently in natural language meaning.

Frege’s *sense* (*Sinn*), as contrasted with *reference* (*Bedeutung*), saved formal semantics from the problem of substitutivity of coreferential expressions. For example, since it is true that Hilary Mantel is the author of *Wolf Hall*, it should be possible to substitute ‘Hilary Mantel’ for ‘the author of *Wolf Hall*’ as in (6), preserving the truth of the sentence.

(5) The author of *Wolf Hall* is visiting Cambridge this spring.

(6) Hilary Mantel is visiting Cambridge this spring.

However, although Hilary Mantel is identical to the author of *Wolf Hall*, sentence (7) differs substantially from (8); (7) is informative while (8) is not.

(7) Hilary Mantel is the author of *Wolf Hall*.

(8) Hilary Mantel is Hilary Mantel.

These two ways of referring to the winner of the 2009 Man Booker Prize come with different ways of thinking, ways of understanding, or, to use the celebrated phrase, different *modes of presentation* of the referent. The senses we grasp when we understand these two expressions are different. Now, when we embed (7) and (8) in the contexts expressing mental attitudes, using intentional verbs such as ‘believe’ or ‘doubt’, Frege’s concept of sense acquires a new importance for semantic theory. In (9) and (10), using Frege’s own explanation, the *sense* of ‘Hilary Mantel’ plays the role of the *referent*.

(9) Harry believes that Hilary Mantel is the author of *Wolf Hall*.

(10) Harry believes that Hilary Mantel is Hilary Mantel.

Analogously, in (11), the sense of the definite description ‘the author of *Wolf Hall*’ plays the role of the referent. The description refers not to its customary reference, but instead to the sense. Sense, however, is more than the speaker’s way of thinking; it is an intersubjective way of thinking. It can function as the speaker’s own mode of presentation of the referent, but equally can it function as someone else’s way of thinking about this referent.

(11) Harry believes that the author of *Wolf Hall* is a man.

This objectivity of sense shifted intentionality from the domain of the individual and the mental to the domain of the external. For Kripke (1972) and Putnam (1975), proper names and natural kind words acquire reference through the causal link with the world. The topic of the externalist/internalist debate, albeit very important, is tangential to the present discussion of intentionality vis-à-vis intentions and will not be pursued here. Instead, we will now focus on one aspect of the debate, namely the role of intentionality in the theory of meaning.

Let us now return to the topic of ‘antipsychologism’ of analytic philosophy. Frege can be credited with developing a new concept of logic. In his *Begriffsschrift* (Frege 1879), he developed an analysis of the logical form of sentences in terms of predicates and arguments, where the reference of a predicate is a function from objects to truth values. This development marked the end of the era of psychological logic that studied thought processes and subjective mental representations. According to Frege, the object of study of logic is not the agent who uses the rules of inference, but the rules of logical inference themselves, and thereby the languages of logic themselves. In *Grundlagen der Arithmetik*, Frege (1884) offers a new, depersonalised stance on truth, definitions, logic, mathematics, and indirectly on natural language semantics. He says that ‘[t]here must be a sharp separation of the psychological from the logical, the subjective from the objective’ (p. 90). The ways of thinking about objects must not be confused with the objects themselves. Instead of focusing on someone’s ‘thinking that something is true or valid’, we now focus on truth and validity as such. To give further examples, in his *Grundgesetze der Arithmetik*, (Frege 1893, vol. 1: 202), he calls the effect of psychology on logic a ‘corrupting intrusion’ because it has to be emphasised that ‘*being true* is quite different from *being held as true*’ (p. 202). In *Logic*, he stresses that ‘Logic is concerned with the laws of truth, not with the laws of holding something to be true, not with the question of how people think, but with the question of how they must think if they are not to miss the truth.’ Frege (1897: 250).

This ban on psychologism has never been lifted since.¹ Frege’s function-argument analysis, juxtaposed with Tarski’s semantic definition of truth, led to the development of formal semantics of natural languages. Now, this rejection of psychology can easily be taken to suggest that the very idea of intentionality had to be poured out with bath water as well.

¹ However, Travis (2006: 125-6) argues that holding any view on how logical laws apply to thinking subjects constitutes a form of psychologism and thereby dissociating subjects from logical laws à la Frege is a form of psychologism as well.

After all, what we have here is Frege's forceful rebuttal of Husserl's stance on logic as a study of mental processes.² As can be seen from the following section, however, intentionality need not necessarily pertain to mental states and proves to be a useful theoretical tool that can easily be dissociated from 'psychologising'.

The follow-up question to ask is whether intentionality is a property of mental states only, or rather it can also be construed as a property of other objects. The answer gleaned from subsequent theorizing is definitely positive. For the current purpose, however, we will not be concerned with the intentionality of human organs and limbs, or intentionality of information-processing systems discussed by Millikan and Dretske (see e.g. Lyons 1995; Jacob 2003). Neither will we be interested in Fodor's (e.g. 1975, 1981) discussion of intentionality as a feature of the brain. What we will focus on is the intentionality of linguistic expressions.

For Searle (1983, 1992b), our beliefs and intentions have intrinsic, basic intentionality, while linguistic expressions have derived intentionality in the sense that the meaning of acts of speech can be analysed in terms of intentional states, such as belief or intention. In other words, Searle says that the mind 'imposes' intentionality, so to speak, on linguistic expressions in that the basic intention to represent is responsible for the derived intention to communicate. The intentionality of the mental state that underlies the act of communication bestows on that act so-called *conditions of satisfaction*. In brief, beliefs have intrinsic intentionality, while utterances have derived intentionality (Searle 1991: 84), or "I impose Intentionality on my utterances by intentionally conferring on them certain conditions of satisfaction which are the conditions of satisfaction of certain psychological states" (Searle 1983: 28).

What Searle proposes here is a so-called *double level of intentionality*. Mental states such as beliefs, wishes, or hopes impose conditions of satisfaction on the expressions of these states. These conditions, in turn, play a major role in determining the meaning of the linguistic expressions. For example, a request may inherit conditions of satisfaction from the speaker's wish or desire. An important question arises at this point, namely what exactly does it mean to impose, or confer, conditions of satisfaction by one object on another? How can one transfer them from a wish, a mental state, to a request, a linguistic object? Harnish (1990:

² Cf. Frege's letters to Husserl (Frege 1906) and his review of Husserl's *Philosophy of Arithmetic I* (Frege 1894).

189) points out that an *intention to confer* them does not suffice because there must be a constraint that one cannot intend and fail. Furthermore, there must be a restriction on types of objects from and to which intentionality can be transferred. As was proposed in Jaszczolt (1999: 106), we should divert our attention from the conferment as such and direct it instead to the fact that *the same* conditions of satisfaction pertain to the mental and to the linguistic. Once we pose the problem in this way, what remains is to take a stance on the relation between the mental and the linguistic, perhaps arguing as follows. Since the double level is supposed to characterise speech acts themselves rather than the relation between a speech act on the one hand and a mental state on the other, then we should think that speech acts have both basic intentionality, *qua* externalizations of mental states, and derived intentionality, *qua* linguistic objects. Since language is one of the vehicles of mental states, this seems to be a natural way of explaining the mysterious and rather unfortunately named ‘double level’ of intentionality.³ By this reasoning, intentionality is always an intrinsic rather than a ‘conferred’ property.

Next, just as intentionality is a property of linguistic acts, so having intentions is a property of their owners. The latter is the topic to which we now turn.

3. Intentions in communication

3.1. Intentions and inferences

Intentions and the study of language communication have long and intertwined traditions. For John Locke, language is there to fulfil the intention of expressing the thoughts of their holders. In a similar vein, for Grice, the concept of meaning is founded on what is communicated, intentionally, by the speaker. Analysing sentence meaning takes us only part of the way; without employing the communicative intention the analysis is incomplete. In order to present the principles on which his theory of meaning is founded, we have to begin with Grice’s seminal paper ‘Meaning’ (Grice 1957). Firstly, he points out the difference between the so-called *natural meaning* and *non-natural meaning*. When meaning that *p*

³ For a detailed discussion see Jaszczolt 1999, chapters 3 and 6.

entails that it is the fact that p , we have an instance of natural meaning. (12) is an example of natural meaning in that the symptom and the disease are linked through a natural connection.

(12) These red spots mean meningitis.

This kind of meaning is of no interest to a linguist. Instead, linguists should focus on *speaker's meaning* or *non-natural meaning*, also known as $\text{meaning}_{\text{nn}}$. $\text{Meaning}_{\text{nn}}$ is conventional, not characterised by the relation of entailment discussed above, and it is this kind of meaning that is the object of study in Gricean and post-Gricean pragmatics. Speaker's intentions are crucial for defining $\text{meaning}_{\text{nn}}$, and so is, to a greater or lesser degree depending on the approach, the recognition of speaker's intentions by the addressee. Grice utilises intentions in the definition as follows:

'A meant_{nn} something by x ' is roughly equivalent to 'A uttered x with the intention of inducing a belief by means of the recognition of this intention'." (Grice 1957: 219),

and elaborates further in 'Utterer's Meaning and Intentions' (Grice 1969: 92):

'U meant something by uttering x ' is true iff [if and only if], for some audience A, U uttered x intending:

[1] A to produce a particular response r

[2] A to think (recognize) that U intends [1]

[3] A to fulfill [1] on the basis of his fulfillment of [2].

This definitional role of intention was also essential in the speech-act literature of that period (cf. Austin 1962; Searle 1969).

The reliance of pragmatic theory on intentions does not mean, however, a return to psychologism addressed in Section 2. Making a definitional use of them need not lead to the inclusion of a theory of mental processes in pragmatics. Further developments in the post-Gricean tradition testify to a free choice here: some pragmaticists stayed close to Grice's spirit and upheld the antipsychological stance (e.g. Levinson 2000), focusing on general principles of rational action, but as a matter of methodological assumption, without an investigation of the psychological processes underlying linguistic communication,⁴ while others ventured into cognitive science and discussions of inferential processes that lead to the

⁴ Cf. Levinson's (2000) word- or phrase-based, and a fortiori language system-based defaults.

intended or recovered utterance meaning (Sperber and Wilson 1986; Recanati 2004a; Jaszczolt 2005). It has to be remembered at this juncture that for Grice the recognition of the speaker's intentions need not always mean conscious and laborious processing. The recovery of the intention can be 'short-circuited', so to speak, when the meaning is conventionalized in a language and the conventions create a 'shortcut' through the recognition of the intentions. It can also be short-circuited when the intended content can be presumed in the particular context. Default meanings of the first type originated in Grice's concept of the generalized conversational implicature and have been developed further in the theory of presumptive meanings (Levinson 2000). Default Semantics (Jaszczolt 2005, 2010c) accounts for both types.⁵ Some examples are given in (13)-(17), with the convention-driven or context-driven defaults in (13a)-(17a).

- (13) Some people like jazz.
- (13a) Not all people like jazz.
- (14) It is possible that soon all cars will run on electricity.
- (14a) It is not certain that soon all cars will run on electricity.
- (15) A secretary brought us coffee.
- (15a) A female secretary brought us coffee.
- (16) A Botticelli was stolen from the Uffizi.
- (16a) A painting by Botticelli was stolen from the Uffizi Gallery in Florence.
- (17) Kate and Leonardo performed superbly in *Revolutionary Road*.
- (17a) Kate Winslett and Leonardo diCaprio performed superbly in *Revolutionary Road*.

All in all, the possibility of such a non-inferential uptake notwithstanding, it is evident that Gricean accounts explain communication in terms of **intentions** and **inferences**. An intention to inform the addressee is fulfilled simply by the recognition of this intention by the addressee. We devote more attention to the types of intentions in communication in Section 3.2, and the range of inferences underlying communication in Section 3.3.

⁵ See also Davis (1998) on the conventionality of sentence implicature.

Meaning_{nn}, explained in terms of intentions and inferences, provided the foundations for Grice's theory of a cooperative conversational behaviour and thereby his theory of implicature. According to this theory, interlocutors are rational agents whose behaviour is governed by the Cooperative Principle: "Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged" (Grice 1975: 26).⁶ Implicatures are normally considered to be meanings intended by the speaker and recovered thanks to this rationality assumption. To use Horn's (2004: 6) apt dictum, 'Speakers implicate, hearers infer'. However, in literary texts, for example in poetry, the writer may intentionally leave the choice of possible implicatures open for the reader. And even in interaction, speakers may be taken to be implying something that is contrary to their intention (so-called 'unintended implicature') (Cummings 2005: 20-21; Haugh 2008b), or may intentionally leave the interpretation of what has been said or implied open to the hearer (Clark 1997; Jaszczolt 1999: 85; Haugh forthcoming). It has to be pointed out that originally, in Grice's writings, the term 'implicature' referred to the process of intentionally conveying some meaning in addition to what is said (and thus was restricted to the speaker), whilst the product of this process was called an 'implicatum' (based on inferences by either the speaker or the hearer). Gradually, however, the term 'implicature' took over to serve both functions, obscuring the distinction between 'utterer-implicature' and 'audience-implicature' (Saul 2002a; see also Horn this volume).

It is evident that implicatures are pragmatic constructs through and through. They pertain to communicated thoughts and need not, or according to some post-Griceans must not, have direct counterparts in uttered sentences. They stand for inferred meanings and one of their most interesting characteristics is their cancellability. For example, we can always cancel the implicature in (15a) as in (15b).

(15b) A secretary brought us coffee. He was smartly dressed and moved swiftly.

This property is particularly interesting in that it clearly allows for gradation. In post-Gricean pragmatics, one of the main bones of contention has been the delimitation of what is said: contextualists, such as Sperber and Wilson (e.g. 1995b) or Recanati (e.g. 1989, 2004a) provide for the development of the logical form of the sentence until it represents the content understood to be the main message communicated by this utterance. Such an extended unit is

⁶ But see also Mill (1872: 517); Ducrot (1972: 134).

not, however, dubbed an implicature; it is an *explicature* or *what is said*, respectively for these authors. Implicatures have to pertain to separate, additional thoughts with their own pragmatic force (Haugh 2002: 128-130). Default Semantics, a more radical contextualist post-Gricean approach, lifts this restriction on what is said by the utterance and allows for the main intended message, or primary meaning, to be independent from the logical form of the uttered sentence. This is where the interesting property of cancellability comes into the picture. When we try to define implicatures *qua* separate thoughts as cancellable meanings, while what is said *qua* enriched, modulated, extended sentence meaning, we soon realise we are barking up a wrong tree and are forced to retreat all the way to Grice (1978). Implicatures can be dubbed cancellable only if we maintain Grice's original all-encompassing definition of conversational implicature,⁷ also pertaining to the cases of the development of the sentence meaning such as (18a).

(18) Sue took a sharp knife and chopped the onions.

(18a) Sue took a sharp knife *and then* chopped the onions.

When we allow for such cases of sentence embellishment to be classified as *what is said* or *what is explicit*, leaving the term *implicature* only for instances of separate thoughts communicated by means of the uttered sentence, then the property of cancellability becomes much less clear-cut than it was on Grice's original account. In theory, what we would want in order to strengthen the rationale for the *what is said/what is implicit* distinction drawn in this way is the property of cancellability to apply to implicatures but not to what is said. This, however, is not the case. It is not the 'enriched said' vs. implicit distinction that marks the boundary between the cancellable and non-cancellable meanings, but rather, when our aim remains Gricean meaning_{nn}, we have to acknowledge that cancellability is a gradable property, dependent on the speaker's intentions. When the explicit content corresponds to the main speaker intention, cancellability is unlikely; likewise, when one of the implicit contents corresponds to the main intended meaning, this implicature is also entrenched. All in all, we have to either remain minimalist about semantics and construe what is said as sentence meaning *tout court* (Borg 2004; Cappelen and Lepore 2005a) and deny it the property of

⁷ Conventional implicatures are not included in the discussion as they have been conclusively demonstrated to be redundant in pragmatic theory. See e.g. Bach (1994b).

cancellability, or we should be contextualist about meaning and tie cancellability to the strength of intending, disregarding the explicit/implicit distinction.⁸

3.2. Types of intentions

The intention on which Grice's theory of meaning_{nn} is founded can be called *communicative* intention: an intention to communicate certain content to the audience. It is fulfilled by its recognition. Bach and Harnish (1979: 7) call it an *illocutionary-communicative* intention and found it on the so-called *communicative presumption*: an assumption that when the speaker utters something to the addressee, the speaker is doing so with an illocutionary intention. Potentially ambiguous utterances result in unambiguous acts of communication thanks to the recognition of the speaker's intentions. Analogously, sentences with indexical terms result in referentially complete utterances because of the recognition of the speaker's referential intention. Referential intention is understood here as part of the overall communicative intention.

The inherent reflexivity of the communicative intention is what, according to Levinson (2006: 87), "makes open-ended communication, communication beyond a small fixed repertoire of signals". There is some debate, however, around whether the communicative intention involves one or two degrees of reflexivity.⁹ Grice's original formulation of communicative intention involved two levels of reflexivity: a first-order intention (to intend to inform or represent something) embedded in a second-order intention (to intend that this first-order intention be recognised by the hearer), which was further embedded in a third-order intention (to intend that the recognition of the first-order intention be based on the hearer recognising the speaker's second-order intention). The utility of this third-order intention, however, has been disputed (Bara 2010: 82-83). Searle (2007: 14), for instance, argues that Grice's third-order intention ties meaning_{nn} to perlocutionary effects thereby confounding meaning with "successful" communication.

Subsequently, communicative intention has thus been reanalysed as only involving two kinds of intentions, the latter embedded in the first: the communicative intention and the informative intention. Communicative intention in this view consists of making it obvious to

⁸ This discussion is developed in detail in Jaszczolt (2008).

⁹ Although see Recanati (1986), and more recently Davis (2008), who question the necessity for communicative intentions to be reflexive at all (cf. Bach 1987; Witek 2009).

the addressee that the speaker has an intention to inform him/her about something (Sperber and Wilson 1986: 61). Informative intention can remain covert when it is not issued as part of the communicative intention. In other words, A may let B know something without letting B know that A wants B to know. Puzzles carried by examples of situations contrived to demonstrate communication through hidden intentions, and implications for the definition of meaning_{nn}, have exercised many philosophers, and it has been proposed that “manipulative intentions” may underlie informative and communicative language use in some instances (Nemeth T. 2008).

Searle’s (1983) work on intentionality introduced a further distinction between prior intentions and intention-in-action, the latter referring to ‘the proximal cause of the physiological chain leading to overt behaviour’ (Ciaramidaro et al 2007: 3106; see also Pacherie 2000: 403). The different types of intention said to underlie communication have subsequently proliferated as intention has been variously used to refer to “action planning and representation, goal-directedness and action control” (Becchio and Bertone 2010: 226), thereby encompassing a continuum from states of mind to actions (Pacherie 2003: 599). In the remainder of this section we concentrate on exploring different types of prior intentions, rather than action intentions, as it is the former that are arguably the most relevant to the analysis of (speaker) meaning in pragmatics.

The notion of prior intention was initially proposed by Searle (1983: 165-166) to encompass the communicative intention (or what he prefers to call meaning intention, encompassing a first-order representation intention and a second-order communication intention). However, subsequent work has indicated that there are a range of different types of prior intentions, including not only communicative and informative intentions, but also future-directed/higher-order intentions, private intentions, and we-intentions. The latter are ontological ambiguous, however, as to whether they are actually, in practice, prior intentions, or better characterised as (post-facto) “emergent” intentions.

The first distinction made in relation to prior intentions is between those which are present-directed (or proximal) and those which are future-directed (or distal) (Bratman 1987, 1999; Ciaramidaro et al 2007: 3106).¹⁰ While the communicative intention is essentially present-directed, being used to account for speaker meanings at the utterance-level, it has

¹⁰ Cf. Fetzer’s (2002) distinction between macro and micro communicative intentions. The former type appears to be largely analogous with future-directed/higher-order intentions.

become apparent that higher-order intentions, controlling “whole segments of dialogue” (Tirassa 1999) or the planning of activity types, including long-term goals (Bratman 1999), may also be relevant to speaker meaning.

The analytical import of considering higher-order intentions is underscored in Ruhi’s (2007) analysis of compliment responses. In the excerpt below, a compliment (you’re a good cook) is deployed by one family member to another in order to imply a request (that the receiver of the compliment cook again for guests).

(19a)

1 Aysun: Ayhan! I didn’t know you were so skilled [at cooking] darling

2 Ayhan: Go on *canım*! You continue thinking I’m inept (Ruhi 2007: 138)¹¹

Ayhan responds to Aysun’s positive assessment of his cooking in line 1 by questioning the sincerity condition (line 2), thereby displaying uptake of Aysun’s communicative intention to compliment him. However, Ayhan’s response to the compliment also “metapresents the higher-order intention of the C[ompliment] as a request” (Ruhi 2007: 139), as by disputing Aysun’s claim that she thinks Ayhan is skilled in cooking, Ayhan forestalls the implication that Aysun wants Ayhan to cook for the guests (ibid: 138-139). In other words, Ayhan undermines the legitimacy of Aysun’s implied request that he cook for guests by challenging one of its preparatory conditions (i.e., he can cook well).

This analysis is confirmed in the subsequent turns, reproduced below, where it becomes obvious that Ayhan was attempting to pre-empt any future requests for him to cook again.

(19b)

3 Aysun: How can that be! From now on we’ll discover your talents
every time we have guests for dinner

4 Ayhan: Don’t exaggerate Aysun! This was just a one-off (Ruhi 2007: 138).

Ruhi goes on to argue that such an example challenges the view that meanings can be analysed solely in terms of (communicative) intentions, and suggests that meanings and

¹¹ The original excerpt was in Turkish, but only Ruhi’s translation is reproduced here. The term *canım* (‘my dear’, lit. ‘my life/soul’) is used here as a mitigator (Ruhi 2007: 113).

actions are “hierarchically controlled by higher-order intentions that affect single speech act production and the unfolding of discourse” (Ruhi 2007: 139).

A second distinction made is that between private intentions (involving the representation of a goal that involves only the speaker) and social intentions (involving the representation of a social goal, where the speaker’s goal involves at least one other person in addition to the speaker). While it might be assumed at first glance that only prior intentions involving the representation of social goals could be relevant to communication, it is evident that private intentions may also become salient in some cases, for example, when inferences about private intentions (e.g. the hearer reaching into her purse to pay for her coffee) may enter into a context salient to formulating a social intention (e.g. the speaker offering to pay for the hearer’s coffee).

Finally, the argument that certain joint, cooperative activities, including communication, cannot be straightforwardly reduced to individual intentions has led to the introduction of the notion of we-intention into this increasingly complex landscape (Searle 1990). One key question around the notion of we-intentions (sometimes also called a shared or joint intention) is whether they can be reduced to individual intentions supplemented with mutual beliefs (Tuomela and Miller 1988; Bratman 1992, 1993, 1999; Pacherie 2007), or whether they constitute a primitive form of intentionality that cannot be reduced to individual intentions (Becchio and Bertone 2004: 126; Searle 1990; Tuomela 2005).

According to Bratman’s (1992, 1993) account, which is defended by Pacherie (2007), shared intention can be explicated in terms of individual intentions in the following way:

Where J is a cooperatively neutral joint-act type, our J-ing is a shared cooperative activity only if:

- (1) (a) (i) I intend that we J.
- (1) (a) (ii) I intend that we J in accordance with and because of meshing subplans of (1) (a) (i) and (1) (b) (i).
- (1) (b) (i) You intend that we J.
- (1) (b) (ii) You intend that we J in accordance with and because of meshing subplans of (1) (a) (i) and (1) (b) (i).

- (1) (c) The intentions in (1) (a) and (1) (b) are minimally cooperatively stable.
- (2) It is common knowledge between us that (1) (Bratman 1992: 338)

While Pacherie (2007) argues that this formulation of shared intention captures the commitment of the participants to a joint activity (conditions (1) (a) (i) and (1) (b) (i)), the mutual responsiveness of each participant to the other (conditions (1) (a) (ii) and (1) (b) (ii)), and commitment to mutual support (condition (1) (c)), Becchio and Bertone (2004: 127-128) argue that a formulation of shared intention building on mutual belief about individual intentions is “cognitively implausible” as the requirement for sharedness is too strenuous.

Searle (1990: 407), on the other hand, defines a we-intention as simply “we intend that we perform act A”, which in turn presupposes *cooperation* on the part of the we-intenders. He argues that “we-intentions cannot be analyzed into sets of I[individual]-intentions, even I-intentions supplemented with beliefs” (p.404), giving the example of playing football to illustrate his point:

Suppose we are on a football team and we are trying to execute a pass play. That is, the team intention, we suppose, is in part expressed by “We are executing a pass play.” But now notice: no individual member of the team has this as the entire content of his intention, for no one can execute a pass play by himself. Each player must make a specific contribution to the overall goal [...] Each member of the team will share in the collective intention, but will have an individual assignment that is derived from the collective but has a different content from the collective. Where the collective’s is “We are doing A,” the individual’s will be “I am doing B,” “I am doing C,” and so on. (Searle 1990: 403)

The essence of Searle’s argument, then, is that each member of the football team I-intending different parts of the action of pass play does not simply add up (i.e. summatively) to the collective action of a pass play, because a level of cooperation is presupposed that goes beyond each person “doing their part” (ibid: 405). Cooperation implies, we would suggest, that each team member’s I-intentions are responsive to their perceptions of the I-intentions of other team members, meaning, in other words, that the I-intentions of team members are both afforded and constrained by the I-intentions of others in making the pass play. In order for these I-intentions to be responsive in this manner, it is claimed that the team members must be we-intending to engage in this collective activity.

This description of a particular collective action bears remarkable similarities to the cooperative nature of communication assumed by Grice, with a number of scholars noting, in passing, the potential importance of we-intentions for the analysis of communication (Becchio and Bertone 2004: 132; Clark 1996, 1997; Gibbs 1999, 2001; Searle 1990: 415). The place of we-intentions in analysing speaker meaning and communication, however, has received only passing attention in pragmatics to date. This is perhaps due to the inherent ontological ambiguity of we-intentions (Haugh 2008c: 53-54): while they are characterised as prior intentions in the minds of speakers and hearers, there is equivocality about how such an a priori mental state comes to be shared between two or more people in the first place (Fitzpatrick 2003; Velleman 1997), even in models which argue for their importance (Tomasello and Rakoczy 2003; Tomasello et al 2005).

Consider, for instance, Levinson's (1983: 358) example of a "transparent pre-request" in line 1 in the excerpt below:

(20)

- 1 A: Hullo I was wondering whether you were intending to go to Popper's
 talk this afternoon
- 2 B: Not today I'm afraid I can't make it to this one
- 3 A: Ah okay
- 4 B: You wanted me to record it didn't you heh!
- 5 A: Yeah heheh
- 6 B: Heheh no I'm sorry about that... (Levinson 1983: 358)

While it is not entirely clear that B was in fact treating A's utterance in line 1 as a pre-request, or simply as a request for information from his response in line 2, it is evident by line 4 that B has understood A's initial question as a pre-request through his topicalisation of A's *intentions* ("wanted me to record it") (Haugh 2009: 95). However, in order for B to infer the communicative intention underpinning A's question in line 1 (i.e. the intention to check the preparatory conditions for a forthcoming request that B record (Karl) Popper's talk), it appears an inference about A's higher-order intention (to launch a request sequence) is also required. This, in turn, presupposes that A and B are we-intending engagement in a particular activity, namely, conversational interaction. The issue is, however, and this proves crucial to the interpretation of A's communicative intention underlying his utterance in line 1, whether

they are we-intending the joint activity of phatic/relational talk (consistent with an interpretation of A's communicative intention simply being a request for information about B's *intentions* in relation to Popper's talk), or they are we-intending the joint activity of goal-oriented talk (consistent with A's higher-order intention of negotiating a request that B record the talk for him).¹² In other words, the communicative intention of A is embedded within his higher-order intention, with both intentions arguably being further embedded within a we-intention.

This requirement for multiple types of prior intentions embedded within each other, however, leads to analytical equivocality on two counts. First, the requirement for inferences directed at both A's communicative intention (which is present-directed) and his higher-order intention (which is future-directed) leads to trouble in ascertaining when exactly this implied request arises. Second, there is equivocality about how this we-intention is shared in the first place. It is not until line 4 in this interaction that A and B could plausibly be taken to be we-intending the joint activity of goal-oriented talk (in the context of which a request sequence makes sense). Searle's characterisation of a we-intention as a prior intention thus strikes very real problems when applied to the analysis of joint activities such as conversation. It appears, then, that what underpins the request implicature here might be better characterised as a kind of "emergent" intention (Clark 1997; Gibbs 1999: 38, 2001; Haugh 2007c: 95, 2008c, 2009; Kecskes 2010a: 60-61).

Finally, it is also worth noting in passing that in example (20) above, A also makes reference to B's (higher-order) intentions ("*intending* to go") in line 1, while A ratifies in line 5 ("yeah") B's preceding topicalisation of his intentions. It appears, then, that the folk notion of *intending*, encompassing (1) the expression of future plans for self, (2) ascribing to or asking of others their future plans, (3) describing what oneself or others want to achieve by doing or saying something, and (4) classifying actions as being done with the speaker's awareness of the implications of them (Gibbs 1999: 22-23), may also be sometimes relevant to the analysis of speaker meaning. The notion of higher-order intention appears to overlap with senses (1) through to (3) of the folk notion of *intention* while the last two senses of *intention* appear consistent with the claim in ethnomethodological conversation analysis that speakers are held (morally) accountable for their meanings (Garfinkel 1967; Sacks

¹² Although this is not to say there are not relational implications arising in the case of goal-oriented talk.

1964[1992: 4-5]; Heritage 1988).¹³ While there has been work focusing on intuitive, folk understandings of *intention* (Breheny 2006; Gibbs 1999, 2001; Knobbe 2003; Knobbe and Burra 2006; Malle, Moses and Baldwin 2001; Malle 2004, 2006),¹⁴ there remains considerable work to be done in comparing folk understandings of *intention* with the various analytical concepts of intention developed in philosophical pragmatics. Drawing a clear distinction between folk and theoretical/analytical notions of intention is clearly important, as it is through their common link to folk notion of *intentional* actions that intentionality (encompassing the directedness/aboutness of linguistic acts) and speaker intentions (encompassing the *a priori* goal-directedness and deliberateness of speaker actions) are often confounded. It is only by making such comparisons that will we better understand the relationship between the intuitive notions of *intention/intending* and the more technical notions we have reviewed here.

3.3. From mental acts to communicative acts and types of inference

The detailed work on different types of (prior) intention and their relationship to speaker meaning has been expanded upon in two quite different ways in pragmatics, as noted in section 3.1. Those more closely aligned with the spirit of Grice's original program have focused on moving from mental acts (intentions) to communicative acts (including both the so-called neo-Griceans and speech act theorists). Those taking a more psychological or cognitive stance, in contrast, have focused on explicating the inferential processes leading from communicative acts to intended or recovered speaker meaning (the so-called post-Griceans).

In Searle's (1983, 2007) work on the relationship between speech acts, speaker meaning, and intention(ality), he claims that for a speech act to be communicated a double level of intentionality is required, as noted in Section 2. This double level of intentionality builds on the notion of direction of fit. Word-to-world direction of fit encompasses the assertive class of speech acts (e.g. statements, assertions etc.), which are "expressions of beliefs and are supposed, like beliefs, to represent how the world is" (Searle 2007: 14), while

¹³ It is this sense of *intention* as a type of moral accountability oriented to by interactants that has been of most interest to those in socio-interactional fields of pragmatics to date (Arundale 2008: 250; Haugh 2008b, 2008c: 69-72; 2009: 108).

¹⁴ Malle and others also make reference to *intentionality* as a folk concept, but as Wierzbicka (2006) points out, intentionality is a technical term that has not actually entered into lay talk.

world-to-word direction of fit encompasses both the directive class of speech acts (e.g. requests, orders, commands, etc.), which are expressions of desires, and the commissive class (e.g. promises, offers, etc.), which are expressions of intention (ibid: 14). Declaratives have a double direction of fit, while expressives (e.g. apologies, thanks, etc.) have a null direction of fit (ibid: 14-15). He gives the example of a speaker wishing to communicate the belief 'It is raining' to illustrate this double level of intentionality:

When the speaker intentionally utters a token of the symbol [It is raining], the production of the token is the condition of satisfaction of his intention to utter it. And when he utters it *meaningfully* he is imposing a further condition of satisfaction on the token uttered. The condition of satisfaction is: That it is raining. The imposition of conditions of satisfaction on conditions of satisfaction is the essence of speaker meaning. (Searle 2007: 23, original emphasis)

According to Searle, then, the first level of intentionality arises from the prior intention of the speaker, namely, that uttering 'It is raining' is a goal-directed and deliberate action. The second level of intentionality arises from the intrinsic aboutness of beliefs. It is by imposing conditions of satisfaction of the belief in question (i.e. that is indeed raining) upon the conditions of satisfaction of the intention in question (i.e. to utter the token 'it is raining') that a speaker can be taken to be meaning this particular assertion. Further work modeling the link between prior intentions (i.e. intentions as mental states) with intention-in-action (i.e. uttering something) has been undertaken (Bara 2010; Pacherie 2006, 2008), in an attempt to move Searle's analytical work into an empirical, testable reality. But such work, albeit very important, is largely tangential to the current discussion since it focuses on investigating neural mechanisms.¹⁵

While Searle makes little reference to inferential work in his analysis,¹⁶ Levinson's (2000) account of implicature is grounded with reference to default logics (for generalised implicatures) and practical reasoning (for particularised implicatures). Default logics aim to capture the notion of reasonable presumption, a *ceteris paribus* assumption. Levinson (2000: 46) argues that default logics capture two important features of generalised implicatures: their

¹⁵ But see also Jaszczolt's (1999: 104-111) criticism of the requirement of the double level of intentionality. Searle (1983) says that the intentionality of linguistic expressions is inherited from the intentionality of the corresponding psychological states. However, instead of the transfer of the conditions from the mental to the linguistic, it is more in line with the intentionality view in philosophy to see language as one of the vehicles of meaning and thereby to regard intentionality pertaining to language as primary rather than derived.

¹⁶ Although see Searle (1975b) for an attempt to model inferences underlying indirect speech acts.

defeasibility and the way in which they involve preferred presumptions. While a number of models of default logics are discussed by Levinson (2000: 46-49), he does not give a definitive answer for the problem of formally modelling the inferences underlying generalised implicatures.

Practical reasoning systems, first mentioned by Aristotle, aim to explicate how speakers reason from a particular goal (i.e. a prior intention) to the means by which this goal is achieved (e.g. an utterance). Levinson argues a modified “Kenny logic”, building on Kenny’s (1966) work on practical reasoning, captures two important features of desirability reasoning: its defeasibility (i.e. context-dependence) and its ampliativeness (i.e. more than one means can achieve an end) (Brown and Levinson 1987: 89). The system formalises inferences that are valid when propositions are “satisfactory” relative to goal wants: “A fiat $F(a)$ is satisfactory relative to a set G of desires $\{F(g_1), F(g_2), \dots, F(g_n)\}$ if and only if, whenever a is true, $(g_1$ and $g_2 \dots$ and $g_n)$ is true.” (ibid: 87). In other words, practical reasoning accounts for how speakers move from formulating intentions (ends) to what is uttered (means). However, they readily acknowledge that a logic-based reasoning system such as this cannot account for how hearers understand speakers’ communicative intentions (ibid: 8).

In moving from communicative acts (utterances) to mental acts (intentions), in contrast, a number of different types of inference have been claimed to underlie speaker meaning, including the recovery of (intended) implicatures. According to some, an implicature is recovered through inductive inference on Grice’s account, and thus constitutes “a probabilistic conclusion derived from a set of premises that include the utterance and such contextual information as appears relevant” (Grundy 2008: 102). It is argued that the probabilistic nature of inductive inferences accounts for mismatches between what the speaker intends and how this is interpreted by the hearer (ibid: 102). Others have claimed that the recovery of conversational implicatures by hearers is better explicated with reference to abductive reasoning (i.e. inference to best explanation) (Allot 2007), or its formal analogue, inference to best interpretation (Atlas 2005; Atlas and Levinson 1981). The key difference between induction and abduction is that the former involves reasoning “from particular instances to a general hypothesis”, while the latter involves reasoning where “the conclusions are based on a best guess” (Allan 2006c: 652).

Somewhat controversially, in their *Relevance* book, relevance theorists Sperber and Wilson have proposed that deductive reasoning is central to utterance interpretation (Sperber

and Wilson [1986]1995b: 69; see also Wilson and Sperber 1986: 45). In the following example, Mary implies that she wouldn't drive a Mercedes in response to Peter's question.

(21) Peter: Would you drive a Mercedes?

Mary: I wouldn't drive *any* expensive car. (Sperber and Wilson 1995: 194)

It is claimed by Sperber and Wilson that this implication can be deduced through the addition of encyclopaedic information about expensive cars ("a Mercedes is an expensive car") as an additional premise together with Mary's assertion that she "wouldn't drive *any* expensive car" (ibid: 194).

One key problem with a reliance on deductive reasoning in explicating how hearers recover speaker meaning, however, as Cummings (1998, 2005) points out, is how to establish the closed set of propositions required for strict deduction (Allan 2006c: 553; Allot 2007: 60). As she argues,

no set of deductive premises is ever fully circumscribed in the sense of containing all the information that is relevant to the comprehension of that utterance – for every set of such premises, some factor that is not part of the set can nonetheless be shown to be integral to the comprehension of that utterance. (Cummings 2005: 130)

In other words, there is no solution given in Relevance Theory as to how the hearer decides which premises to include and those to be excluded, except by the arguably circular argument that the premises selected are determined by calculations of relevance,¹⁷ which in turn determines what the hearer understands to have been implied by the speaker. The circularity of their approach as presented in that book lies in the fact that calculations of the relevance of the contextual premises are determined by the hearer's calculations of the relevance of what is implied.¹⁸

As Cummings (2005: 108) has argued, there remains considerable work to be done to clarify the types of inference that underlie the recovery of speaker meaning. In line with the emphasis on non-deductive inference present in the pragmatic literature in the last decade (see e.g. Levinson 2000, or, in a different paradigm, Asher and Lascarides 2003), she discusses two further types of inference that are crucial to this process: elaborative inferences

¹⁷ Relevance is roughly defined as positive cognitive effects relative to processing effort (Sperber and Wilson 1995b: 265-266).

¹⁸ But see also later discussions of inference, e.g., Carston (2002, 2007).

and presumptive reasoning. Elaborative inferences are knowledge-based inferences used to “establish causal connections between events and construct intentional relations between actions within reasoning” (Cummings 2005: 91), with this knowledge often being of “behaviour tendencies and everyday routines” that serve “to specify normality or typicality conditions on inference” (ibid: 93). She claims that the implicature arising in the example below, for instance, cannot be recovered without elaborative inference.

(22) A: I’m out of milk.

B: There’s a shop at the end of the street. (Cummings 2005: 102)

In order for A to understand that B is implying that A can get milk from the shop at the end of the street, real-world knowledge is required: that corner shops stock commonly used groceries, including milk (ibid: 103). She goes on to claim that “knowledge provides a cohesive link between the utterances of the above exchange – it is through our knowledge of corner shops and their merchandise that we are able to establish the relevance of B’s utterance to A’s utterance in this exchange” (ibid: 103). Such elaborative inferences allow hearers to go from “abstract communication norms and principles” to particular communicated meanings (ibid: 104).

Cummings (2005) also suggests that presumptive reasoning is crucial to the recovery of implicatures, since they are always subject to revision based on the addition of further contextual information, but are nevertheless held to be communicated, at least provisionally by hearers. One kind of presumptive reasoning system proposed by Cummings is “argument from ignorance”, where a proposition is accepted as true because there is no evidence that it is not true (or vice versa) (p.109). Thus, while earlier attempts to model the move from communicative acts (utterances) to mental acts (intentions) have tended to rely on traditional forms of inference (deductive, inductive, abductive), strong arguments have been mounted for more attention to be paid to the non-monotonic basis of inference in pragmatics (Cummings 2005: 242; Levinson 2000: 46).¹⁹

However, one problem facing current models of inference that privilege the speaker’s intention in determining whether communication has occurred (whether in moving from speaker intentions to communicative acts or vice-versa), is the failure to “address how the

¹⁹ Non-monotonicity refers to a “presently reasonable inference whose reasonability may be lost upon the admittance of new information” (Woods 2010: 219).

participants themselves could come to know whether the recipient's inference and attribution regarding that intention is to any extent consistent with it" (Arundale 2008: 241). In other words, there is no account in current intention-based models given as to how speakers and hearers determine something has indeed been communicated. Approaches that conceptualise the inferential work underlying meaning and communication as contingent and non-summative (Arundale and Good 2002; see also Haugh 2009) thus arguably deserve further consideration as well.

In the final section we consider the question of where intentions (and inferences about them) can be located, in particular, whether or not they should be analysed as a purely cognitive phenomenon. While the emphasis in the discussion thus far has largely been on speaker intentions underlying utterances, in broadening our scope to consider their place in interaction and broader society, it is argued that intentions are not only traceable to the mind, but also in interaction, and to broader societal norms and discourses.

3.4. Locating intentions

According to the received view of meaning as the intended expression of thoughts, intentions pertain to the mental states of speakers. Recent work in social neuroscience has begun exploring the neural correlates of our ability to attribute mental states to others, including intentions (Walter et al 2004: 1854). This capacity to attribute mental states is termed Theory of Mind (ToM), with it being assumed that without ToM "other people's behaviour would be meaningless from a third person perspective: behaviour would be observed, but the meaning of actions would not be understood" (Ciaramidaro *et al.* 2007: 3111). Experiments combining brain imaging with various kinds of language use prompts have found that a distributed neural system underlies the ToM mechanism, including the right and left temporo-parietal junctions (right TPJ and left TPJ), the precuneus, and the medial prefrontal cortex (MPFC) (Ciaramidaro et al 2007: 3105; Enrici *et al.* in press; Walter *et al.* 2004: 1854). The detection of agency, for instance, has been tied to neural activity in the superior temporal sulcus (STS) (Frith and Frith 2003), while the representation of our own and other people's mental states (including intentions) has been tied to activity in the medial prefrontal cortex (MPFC) (Walter et al 2004: 1854).

Further work on intentions specifically has established that different parts of this distributed neural network are activated depending on the kind of prior intention involved. This lends some empirical support to the hitherto conceptual distinctions discussed in section 3.2, including that between private and social intentions, and within the category of social intentions, between communicative intentions and prospective social intentions, the latter being a form of higher-order intention (Ciaramidaro et al 2007: 3105; see also Saxe 2006; Walter *et al.* 2004), and also between the representation of individual and we-intentions (Becchio and Bertone 2004: 132).

Attempts to model the understanding of intentions (Bara 2010; Pacherie 2006, 2008) have also been given empirical support. Becchio, Adenzato and Bara (2006), for instance, have argued that the recognition, attribution and representation of action intentions can be traced to different forms of neural activity. They hypothesise that the recognition of the intentions of others is partly based on the same areas of the brain that are activated when one (intentionally) performs actions oneself. The neural basis of this has been argued to be “mirror neurons” (i.e. neurons that fire during action execution and action observation) in the premotor cortex (Fogassi et al 2005; Rizzolatti and Craighero 2004; Rizzolatti and Maddalena Fabbri 2007), although mirror neurons do not in themselves provide the basis for different forms of agentive understanding and shared intentionality (Pacherie and Dokic 2006).²⁰ The attribution of intention to a particular agent is traced in Becchio, Adenzato and Bara’s (2006) experiments to another area of the brain, the inferior parietal lobe, while the representation of prior intentions, particularly those which are social, have been traced to the anterior paracingulate cortex (Walter et al 2004). Recent work has also shown that a common neural network is employed when subjects are comprehending communicative intentions whether the prompt is linguistic or gestural (Enrici *et al.* in press). Despite being in its relative infancy, then, social neuroscientists have begun tracing hypothesised mental states (i.e. intentions and inferences about intentions) to specific neural activity, lending support to the kinds of distinctions made in pragmatics between different kinds of intentions.

However, work in psycholinguistics on the comprehension and attribution of intentions suggests that while these may indeed have neural correlates, speakers consistently over-estimate their ability to project intended meanings to addressees (Keysar 1994b, 2000;

²⁰ The importance of mirror neurons to action understanding, however, has recently come under challenge (Hickok 2009, 2010), with the view that they may actually constitute a byproduct of sensorimotor learning being forwarded (Hickok and Hauser 2010; Heyes 2010).

Keysar and Henley 2002). Moreover, hearers often do not routinely consider what the speaker knows (i.e. common ground) or other mental states in interpreting what has been said (Keysar 2007, 2008). This fundamental egocentrism in the early stages of processing language suggests that due caution should be given to interpreting results of experiments attesting to neural correlates of intentions. While no one would suggest that speakers do not at times have intentions motivating them to say things, or that hearers do not at times make conscious inferences about such intentions, the question is the extent to which such an explanation is sufficient to account for speaker meaning and communication (Haugh 2008c: 52, 2009: 93). The evidence from psycholinguistics suggests that while a (prior) intention-based account of speaker meaning may be necessary, it may not be sufficient.

Work from an interactional perspective also attests to the difficulty of locating intentions relative to meanings in discourse (rather than simply relative to utterances). Haugh (2008c, 2009), for instance, argues that the intentions hypothesised to underlie implicatures are temporally, ontologically and epistemologically ambiguous when the analyst attempts to trace them in actual interactional data. In more closely tracing intentions in conversational interaction it becomes apparent that intentions can be characterised as being “emergent”, as both the speaker and the hearer jointly co-construct understandings of what is meant (Clark 1997; Gibbs 1999: 38, 2001; Haugh 2007c: 95, 2008c, 2009; Kecskes 2010a: 60-61). Kecskes (2010a), for instance, argues that John’s initial intention to give Peter a chance to talk about his trip is not realised in the excerpt below.

- (23) John: Want to talk about your trip?
- Peter: I don’t know. If you have questions...
- John: OK, but you should tell me...
- Peter: Wait, you want to hear about Irene?
- John: Well, what about her?
- Peter: She is fine. She has...well...put on some weight, though

(Kecskes 2010a: 60).

Kecskes suggests that John’s original intention is sidelined by Peter talking about Irene, perhaps because he thinks John might want to know about her (being his former girlfriend).

He argues that “it was the conversational flow that led to this point, at which there appears a kind of *emergent, co-constructed intention*” (Kecskes 2010a: 61, original emphasis).

These “emergent” intentions have interesting parallels with the notion of we-intention discussed in section 3.2. However, Haugh (2008c) argues that the relatively static characterisation of we-intentions does not do full justice to the contingent and emergent nature of the inferential work underlying joint, cooperative activities, including conversational interaction. The key difference between the notions of we-intention and emergent intention is that the former assumes a monadic view of cognition, involving “the summative sequence of individual cognitive activities” (Arundale and Good 2002: 124), while the latter builds upon a dyadic view of cognition:

Each participant’s cognitive processes in using language involve concurrent operations temporally extended both forward in time in anticipation or projection, and backwards in time in hindsight or retroactive assessing of what has already transpired. As participants interact, these concurrent cognitive activities become fully interdependent or dyadic. (Arundale and Good 2002: 122)

Yet while there are differences in their underlying frameworks, the view in philosophical and cognitive pragmatics that we-intentions cannot be reduced to individual intentions due to the sharedness requirement being cognitively implausible (Becchio and Bertone 2004: 127-128), and the complementary view in interactional pragmatics that models of individual intentions are “formally incapable of explaining the non-summative effects or emergent properties observable when individuals are engaged in interaction” (Arundale 2008: 243), both strongly suggest that meaning cannot necessarily always be traced analytically to prior intentions of individual speakers (including their communicative intentions). Instead, the interpretation of meaning is “doubly-dynamic” in the sense that it is “created in-between the interlocutors and the hearer may also be given freedom to create assumptions rather than recover them” (Jaszczolt 1999: 76; see also Hirsch 2010). Gauker (2001, 2003, 2008) has also argued that the role of situational inferences has been vastly underplayed in communicative intention-based models of meaning and communication.

However, such views should not be simplistically interpreted as constituting an anti-intentionalism stance as some have recently assumed (Åkerman 2009; Montminy 2010). Instead, there is growing evidence that we need to make clearer distinctions between speaker (intended) meaning which pertains to the subjective processing domain and the utterance

level, and “joint meaning”, which pertains to the interpersonal domain at the discourse level (Carassa and Colombetti 2009; Kasper 2006; Kecskes 2008, 2010a; Kriempardis 2009: 186). Different types of intentions arguably have different roles to play relative to these different types of meaning. Instead of remaining committed to a hardline intentionalist or anti-intentionalist stance, greater dialogue between those with different views on intention lies at the heart of advancing our understanding of meaning and communication (Haugh 2008a).

We have discussed thus far how intentions can be located both in the minds of individuals, as well as more diffusely in interaction, where they emerge “*between* the mind and the world” (Jaszczolt 1999: 117). There is, however, an additional level at which intentions can be productively located, namely, the social (or societal) level of analysis. The focus here is on deontological aspects of intention and intentionality. Philosophers have conceptualised this as commitment to undertake we-intended actions (Gilbert 2009), for instance, or commitment to the truth conditions of utterances (Searle 2007: 33-34). The notion of speaker accountability in ethnomethodological conversation analysis (Garfinkel 1967; Sacks [1964]1992: 4-5), where interlocutors hold themselves and others normatively accountable for the meanings that arise from what is said (Heritage 1984), can also be productively explored in regards to both speaker intentions and intentionality (Arundale 2008; Edwards 2006, 2008; Haugh 2008c, 2009). For instance, the ways in which speakers are held accountable to meanings through topicalising *intentions*, and how this intersects with interpretative and sociocultural norms can be explored (Haugh 2008b, 2008c).

Holding speakers accountable for what they are understood by others to have implied can also enter into broader societal debates, as argued by Haugh (2008b) in an analysis of the discursive dispute arising in the Australian media as to what was intended by comments in regards to the status of women made by a Muslim cleric.²¹ In the following excerpt, the cleric is being interviewed in the controversial wake of the publication of his sermon.

(24) (‘Defending the faith: Sheik Taj Aldin Alhilali’, Australian *60 Minutes*, Channel 9, broadcast 12 November 2006)

1 M: But you’re the grand mufti, you’re the grand mufti. Why would you

²¹ The comments were reported as follows: “If you take out uncovered meat and place it outside on the street, or in the garden or in the park, or in the backyard without a cover, and the cats come and eat it ... whose fault is it, the cats or the uncovered meat? The uncovered meat is the problem. If she was in her room, in her home, in her hijab, no problem would have occurred” (‘Muslim leader blames women for sex attacks’, Richard Kerbaj, *The Australian*, 26 October 2006).

- 2 say something like that which is going to offend everybody?
- 3 H: ((translated)) I say straightaway what is in my heart. I say it.
- 4 M: I hear what you are saying Sheik Alhilali. But, but, you can't
- 5 say these things, and then say I was misunderstood, misinterpreted,
- 6 I meant something else. If you say them they exist. If you say these
- 7 things about rape, about Jews, about militants and bombers if you say
- 8 them people believe you.
- 9 H: ((translated)) My words, as correctly understood, I stand behind.

(Haugh 2008b: 213)

Martin (the interviewer) first questions in lines 1-2 why Hilali would say something offensive when he holds a position of responsibility in the Australian Muslim community (the grand mufti), to which Hilali responds by claiming he always speaks honestly, implying that he does not necessarily always say what others want him to say (line 3). Martin then invokes the folk view of meaning as residing in words (Bilmes 1986), in arguing that Hilali is responsible for how people understand his words (lines 6-8), and that one cannot be absolved from this responsibility by claiming one was “misunderstood” or “misinterpreted” (lines 4-6). This stance, however, is implicitly rejected by Hilali in line 9, when he claims he stands behind the “correct” interpretation of what he implied, in this case, what he *intended* by his comments. The deontological dimensions of speaker meaning were thus clearly topicalised in this particular interview.

In considering the question of where intentions can be located, then, it has become apparent that this concept is deployed in pragmatics for a number of different analytical purposes. In philosophical pragmatics it is used in accounting for how speakers mean things (communicative intentions) or undertake joint activities (we-intentions), although there are varying levels of commitment to the psychological reality of those intentions (Jaszczolt 1999). In cognitive pragmatics there is a more clear commitment to the assumption that the recognition and attribution of (communicative) intentions underlies communication, but there is also consideration of much more finely-grained range of different types of intentions (including the distinction between prior intentions versus intentions-in-action), with work

attempting to correlate neural activity with such distinctions. Here, intentions tend to be conceptualised as being firmly located in the minds of speakers. In contrast, in interactional pragmatics the focus is on examining the relationship between speaker (intended) meaning and joint or interactionally achieved meanings, with the notion of “emergent intention” (Kecskes 2010a) or “emergent intentionality” (Haugh 2008c, 2009) sometimes being deployed to account for the latter. Here, intentions and inferences are treated as contingent and non-summative, arising in the course of interaction, and thus better traced with reference to dyadic views of cognition (Arundale and Good 2002) rather than individual minds. Finally, in more discursive approaches to pragmatics, the analytical focus is on the normative work *intention* does when deployed in discourse or interaction, with a particular emphasis on how speaker commitment or accountability can be disputed. In this approach, intentions can be found to be diffused across social networks, ranging from dyadic units through to larger social groups.

4. Concluding remarks: intentions as ‘creatures of darkness’ or a useful tool?

We are now in a position to address the methodological question as to how pragmatic theory, aspiring to high predictive power, can be founded on intentions and intentionality – the theoretical notions which are inherently so imprecise and, moreover, possibly not directly empirically testable. In other words, this lack of testability may prove not to be a fleeting state of affairs but an inherent property of intentions. In answering this question one has to point out that the advantages seem to outweigh the shortcomings. We attempt to list here a few arguments in favour of an outlook that maintains the importance of intentions for theorising in pragmatics.

[1] If we ban intentions from pragmatics, we have to use a substitute theoretical tool such as default rules of inference, semanticized pragmatic relations between sentences as in dynamic approaches to meaning, or other similar solutions, e.g. constraints of optimality-theory pragmatics. None of these alternative tools has comparable predictive power as far as speaker meaning is concerned. Instead, we are forced to change the object of study of pragmatics from, so to speak, speaker meaning ‘whatever means the speaker may have used to convey it’ to speaker meaning ‘modelled on the fairly probable semantic patterns’,²²

²² Mid-way solutions are possible though: see Default Semantics (Jaszczolt 2005, 2010c)

[2] In the current state of experimenting in neuroscience, it seems very unlikely that intentions can remain creatures of darkness. Instead, they are being correlated with neural activity as discussed in the preceding section. Intentions in communication derive their theoretical status from intentionality of consciousness. The more we know about intentionality in the brain, the more will we know about intentions. The structure of the explanation is already there in the form of Gricean pragmatics; the scientific flesh is being provided as cognitive science progresses.

[3] Language is a vehicle of thought and pragmatic theory of its use in communication should derive from theories of thought. In order to theorise expression meaning (word/sentence meaning), the basic intentionality of thought needs to be taken into account. In this way, the extent to which expression meaning can be productively defined in terms of speaker meaning (intentions), as originally proposed by Grice, may be further explored.²³

[4] The notion of intention (and indeed intentionality) is already being productively deployed in many different ways in pragmatics. While this proliferation can at times create analytical confusion, it is also no doubt reflective of the metaphorical power of intentions and intentionality in advancing our understanding of how speakers mean things through the use of language.

We suggest, therefore, that while they may be difficult to pin down, it is clear that disciplines do not advance by avoiding slippery questions, particularly, when they lie at their very foundations, as do the concepts of intention and intentionality in pragmatics. Ultimately, it is only through refining or even discarding certain views and developing alternatives that we will continue to advance in our theorisation and analysis of meaning and communication.

²³ See Davis (2003, 2008) for further discussion of this issue.